What does the public associate with "climate change"? A study of open-ended survey questions from 2013 to 2021

Endre Tvinnereim, Gisle Andersen, Gisela Böhm & Åsta Dyrnes Nordø

1 Introduction

Climate change is a global challenge that increases as carbon dioxide (CO_2) and other greenhouse gases accumulate in the atmosphere. The 2022 report of the Mitigation working group of the Intergovernmental Panel on Climate Change (IPCC) underscores that, as the global window for avoiding dangerous climate change is shrinking, costs of renewable energy technology are falling, while the need for climate action becomes increasingly urgent and the required actions more drastic. Against this background, citizens across the globe, and notably the young, are increasingly calling for additional action (Pickard, 2022). Yet while popular consent and engagements are necessary components in most mitigation strategies, has popular engagement changed qualitatively over the past decade? Furthermore do young people display different perceptions of climate change than older segments of the population? Surveys based on numerical and fixed-response scales show some differences across these groups. As climate change is a multifaceted issue, the thought processes behind survey responses are less well known, which argues in favor of employing open-ended survey questions with textual responses (Stoneman, Sturgis, & Allum, 2012). However, to our knowledge, no studies on climate change perceptions examine responses to open-ended survey questions over extended spans of time. How we cut emissions and adapt society to climate change will have great significance for the future. To a considerable extent, this can affect the lives of those who are young today. Survey research shows that age correlates with attitudes to climate change. Poortinga, Whitmarsh, Steg, Böhm, and Fisher (2019) summarise the main points well: Older people are more likely to express trend and attribution scepticism and are less likely to express concern about the consequences of climate change. Different approaches to understanding this include the fact that the elderly are more integrated into existing social structures than the young and may be less inclined to accept radical changes. The elderly have also been found to be more value-conservative than the young (Cornelis, 2009, as cited in Poortinga et al., 2019).

Based on the European Social Survey data, Poortinga et al. (2019) show that the age effects exist across different countries, but the effects appear to be particularly strong in northern European countries. In Northern Europe, the authors also find an age effect linked to perceived impacts and concern about climate change (ibid.:32). Data from the Norwegian Citizen Panel support the notion that such a pattern exists in Norway (Gregersen, 2022). These data show that young people are more worried about climate change than older people. In addition, they also indicate

TEXT AND DISCOURSE ACROSS DISCIPLINES: IN HONOUR OF KJERSTI FLØTTUM. Edited by Øyvind Gjerstad and Anje Müller Gjesdal. BeLLS Vol. 13 (2023), DOI 10.15845/bells.v13i1.3855.

that the share of young (born in 1990 or later) who say that they are very worried ("*i svært stor grad*") about climate change has increased since 2015, while the picture is more stable for the other age groups. Thus, there is indication of an age group gap among those most worried about climate change. A natural follow-up question then is if and how we can comprehend these gaps in attitudes between age groups focusing on how they relate to and think about climate change.

Climate change perceptions also show clear gender effects. In general, women tend to worry more about climate change than men (McCright, 2010). At the same time, previous research using open-ended questions on climate change has shown that women tend to emphasize discussions of weather more than men, whereas men talk more about climate change attribution (Tvinnereim and Fløttum, 2015).

In this paper, we present the results of six surveys of the Norwegian Citizen Panel (NCP; https://www.uib.no/en/citizen) using the same open-ended question from 2013 to 2021. We analyze word choices using keyness analysis and changes over time in topic prevalence using topic modeling. In so doing, we apply the framework developed by Tvinnereim and Fløttum (2015) for Wave 1 of the NCP (2013) and extend it to the subsequent waves. We find marked differences in the words used by respondents born after 1990, as they emphasize action-related terms, whereas older respondents use relatively more descriptive words. Many of the topic modeling results from earlier work replicate in the current analysis, but trends over time are not very strong, suggesting continuities in what people associate with climate change.

We define *associations* as the words that emerge from a process where individuals are asked to describe what comes to mind when hearing or reading about a given idea or topic. The concept is elsewhere referred to as *image associations* (Lorenzoni et al., 2006). We access individuals' associations through open-ended survey questions, which have the advantage over closed-ended questions of being more spontaneous and less guided by researcher frames of reference.

Our group has earlier run quantitative text analysis on questions about climate change in Norway (Tvinnereim & Fløttum, 2015; Tvinnereim, Fløttum, Gjerstad, Johannesson, & Nordø, 2017; Tvinnereim, Lægreid, & Fløttum, 2020). Notably, in 2015, we found that associations with climate change may be usefully categorized using structural topic modeling (STM) into four topics: Weather/ice, Attribution, Future/impact, and Money/consumption. We also found several notable correlations between topic prevalence and background variables. For example, older individuals were more likely to mention weather and ice whereas younger respondents focused more on the future. Women were also more likely to mention weather and ice than men, whereas those less worried about climate change are more likely to discuss causes of climate change in their open-ended responses. By contrast, those worried about climate change were more likely to use words related to the Future/impact topic.

Given that almost a decade has passed since the first open-ended questions about climate change from the NCP were analyzed, it is worth examining the trends in public perceptions of climate change that may be measured using this method. In this article we therefore ask what changes have taken place in people's associations with "climate change" over time and whether variations across demographic (age, gender) and attitudinal variables (worry about climate change) are predictable.

2 Data and methods

2.1 Data

The data were collected using the Norwegian Citizen Panel (NCP), a web-based survey infrastructure with more than 10,000 active participants that are sampled from the Norwegian population registry, so that every Norwegian resident above the age of 18 has an equal probability of being invited to join. The sample is thus close to representative of the adult population. The NCP is a collaboration between the University of Bergen (UiB) and NORCE Norwegian Research Centre, and a component of the Digital Social Science Core Facility (DIGSSCORE) at UiB. The NCP is exclusively used for research purposes.

Six waves of the Norwegian Citizen Panel are included in the study: waves 1, 3, 6, 8, 13, and 21 (Ivarsflaten et al., 2013-2021). The data range from 2013 to 2021. The exact dates of the filed work for these waves are as follows: wave 1: 6 November – 5 January 2013; wave 3: 13 October - 27 November 2014, wave 6: 1 March - 19 March 2016; wave 8: 6 March - 9 April 2017; wave 13: October 17 - November 5, 2018; wave 21: 26 May - 15 June 2021. A methodology report of the data collection can be found in Skjervheim et al. (2013 - 2021).

The textual–responses were elicited with the following question wording:

> What comes to mind when you hear the words “climate change”?

In addition, respondent worry about climate change was measured using the wording “How worried are you about climate change?” and five response options ranging from “Not at all worried” to “Very worried.” Data on worry about climate change are not available alongside textual responses in all waves or for all individuals. To compensate for this deficiency, we use data for the same individual respondents but for earlier waves where available.[1] Where earlier waves did not contain such data, the observation was dropped.

Using earlier data in this way means that we need to postulate that worry about climate change is a relatively stable property of each individual respondent, and that worry about climate change at time *t-1* or *t-2* constitutes a good predictor of worry at time *t*. It also means that we will not emphasize changes in levels of worry over time. Using this method, we are able to include a measure of worry about climate change in 11,636 (90.7%) of the observations. While simultaneous measures of worry would have been better, we feel our approach using earlier data points rather than omitting observations is justified because levels of worry do not show significant aggregate changes over time.

Questions in the Norwegian Citizen Panel are never mandatory, opening for missing values. Data on gender and age cohort were derived from the national population registry, and are thus never missing. In total, our data set starts with 12,833 answers to the open-ended question, all with associated data on age, gender, and survey wave.

[1] Specifically, regarding this question, we use responses from Wave 10 to cover for the lack of data in Wave 13, and data from Wave 20 to cover for the lack of responses in Wave 21. To reduce listwise deletion in regression analyses using worry as an explanatory variable, we seek to fill missing values in each wave with values on the same question in the most recent previous wave, where available. This reduces missingness on this variable to 1,155, yielding 11,071 valid responses.

2.2 Methods

We used word frequencies, keyness analysis, and structural topic modeling to analyze our data. All the methods apply a "bag of words" approach, whereby words in a given text or group of texts are registered without regard to word order. In keyness analysis, textual responses were grouped by the contrast groups of interest, such as age or wave. In structural topic modeling, the individual response to the open-ended survey question are treated as the text to be analyzed.

Before the quantitative text analysis, we pre-processed the text data (Grimmer & Stewart, 2013) through stemming, stop word removal, and removal of infrequent terms. Stemming involves combining different grammatical forms of the same word into one, for example, "houses" and "housed" into "house". Stop words are typically short, frequent, and topic-unspecific words such as articles ("a", "the"), pronouns ("I", "they"), conjunctions ("and", "for"), and prepositions ("on", "under"). We also removed words that occured fewer than ten times. Responses rendered empty by pre-processing (for example those containing only words of low frequency or stop words) were dropped from the data set. We thus ended up with 12,226 responses after pre-processing, having removed 607 in pre-processing. The resulting vocabulary had 1,107 unique terms and there were a total of 66,861 tokens in the corpus, not counting repeated tokens within individual responses. The average length of each response was 5.5 tokens, again not counting repeated terms.

Keyness analysis

Keyness analysis compares pairs of texts to identify the words that make up the strongest contrasts between them. Contrasts are sought between the cohort of young respondents (defined as those born 1990 or after) versus the remainder of the sample, women versus men, early versus late waves, and respondents expressing high versus low levels of worry about climate change. Keyness statistics express the degree to which a word or other feature discriminates between two bags of words. We follow Gabrielatos (2018) and use Bayesian information criteria (BIC)[2] and the normalized percentage difference in frequency (%DIFF)[3] to identify the most discriminating words. This means that both frequency and the size of the ratio between the group count towards the significance score. We only include words where BIC > 6, this can be interpreted as strong evidence against the null hypothesis of no difference (Gabrielatos, 2018:240). The %DIFF statistic is an effect-size measure, where positive percentages signify overuse and negative values underuse. We also report the normalized word frequencies (per 1,000 words) in the corpora compared. The `Quanteda` and `KeynessMeasures` packages in R (Benoit et al., 2018) are employed to perform keyness analyses in this paper.

[2] The Bayesian information criterion is used because p-values are not suited to estimate the probability of an observed frequency difference being due to chance (Gabrielatos, 2018: 240). For this calculation we use the log likelihood value of the frequency difference (LL) and the combined size of the two corpora compared (N). BIC = LL – log (N)

[3] Percentage difference uses the normalised frequency of a word (NFC) in the two corpora. A value of zero indicates no difference, a value of 100 indicates twice the frequency in the first corpus, a negative value indicate underuse in the first corpus. Based on (Gabrielatos, 2018), percentage differences are calculated as follows:
% DIFF = (NFC1 – NFC2) *100 / NFC2. When NFC2 is zero the number becomes very large, we show this as «n.a».

Structural topic modeling (STM)
STM (Roberts, Stewart, Tingley, et al., 2014) is an unsupervised method that clusters texts into topics based on word frequencies with the help of numerical or categorical data associated with the texts. We treat individual survey responses as the texts. Each text may belong to several topics. The parameter *topic prevalence* (used interchangeably with *topic proportion*) expresses the degree to which each text belongs to a topic. For each text, the sum of topic prevalences across all topics equals one. We use the in-built linear regression analysis option (R package `stm`) to examine the correlation between topic prevalence and four explanatory variables: wave, age, gender, and worry about climate change. We also run further statistical analyses of the topic prevalence parameters to examine particular subsets of the data.

One recurring issue when using STM and similar clustering methods is how to set the number of topics. We combine quantitative measures of semantic coherence and topic exclusivity with qualitative readings of several model runs to select our topic number of six, see Appendix for details. We select one model using spectral initialization, which yields a stable result.

The approach is similar to the one described in Tvinnereim and Fløttum (2015), and differs mainly by using spectral initialization, having a higher number of topics, employing six rather than one wave of data, and by adding survey wave as an explanatory variable.

The regression model with topic prevalence of one of the topics as the dependent variable takes the following overall form:

$$\text{Topic prevalence} \sim a + b_1*\text{age} + b_2*\text{wave} + b_3*\text{gender} + b_4*\text{worry} + \text{error}$$

We display the effects of each individual variable by holding the other variables constant at their reference categories (not young, man) or at their means (for wave). In cases where we do not expect a linear relationship between topic prevalence and explanatory variables, we simply calculate means and standard errors for the topic prevalence parameter (in STM denoted as "theta") for each category of the explanatory variable.

3 Results

Table 1 displays the word frequencies. The most frequent words in the entire corpus of study are "climate", "human-made, and "more"; followed by "global", "extreme weather", "warming", "weather" "change", and "warmer". Thus, the most frequent words refer to facets of the direct meaning of the word climate change, to the human causation of climate change, and to weather.

Table 1. Word frequency pooled across six waves of open-ended questions about climate change. Data from Norwegian Citizen Panel. Frequency shows the overall word frequencies for the entire corpus.

Rank	Word	Translation	Frequency
1	Klima	climate	1485
2	Menneskeskapt	human-made	1376
3	Mer	more	1278

4	Global	global	895
5	Ekstremvær	extreme weather	892
6	Oppvarm	warming	876
7	Vær	weather	870
8	Endring	change (noun)	856
9	Varmer	warmer	821
10	Mye	much	673
11	Endr	change (verb)	667
12	Klimaendring	climate change	613
13	Temperatur	temperature	598
14	Utslipp	emissions	575
15	Mennesk	human being	529
16	Tenker	think	514
17	Verden	world	491
18	Gjøre	do	490
19	Naturlig	natural	467
20	Store	large	454

3.1 Keyness analysis

The keyness analysis helps us identify the words that are used significantly more or significantly less frequently by a specific group of respondents.

Table 2 shows the keyness analysis by age cohort. Respondents born in 1990 or later are compared to older respondents. Top words distinguishing the younger cohort are “should”, “comes”, and “problem”. They also use words like “think”, “do” “global”, and “warming” more than older cohorts. The young are also characterized by their low usage of words that are related to weather, like “wind,” “wetter,” and “bad weather” (one word in Norwegian).

Table 2. Keyness analysis by age cohort. Respondents born in 1990 or later are compared to older respondents. Relative frequency of a word in the two groups and percent difference in use (%DIFF).

Word	Translation	Rel. freq. young	Rel. freq. other age cohorts	%DIFF
Burd	should	3.9	0.8	363.8
Kommer	comes	7.2	2.8	157.1
Problem	problem	6.7	3.0	126.8
Tenker	think	14.8	6.9	115.2
Gjøre	do	12.5	6.7	85.4
Global	global	19.8	12.6	56.6
Oppvarm	warming	18.8	12.4	51.0
Regn	rain	1.6	5.0	-68.4
Kloden	globe	1.7	5.6	-69.1

Nedbør	precipitation	1.1	5.9	-80.7
Uvær	bad weather	0.7	4.0	-81.9
Våtere	wetter	0.4	2.9	-85.2
Mere	more	0.1	2.5	-94.3
Vind	wind	0.1	4.6	-96.9
Milder	milder	0.0	2.2	-100.0

Table 3 shows the contrast between the first two waves conducted in 2013 and 2014 and responses from wave 13 (2018) and 21 (2021). This comparison allows us to identify significant temporal changes in the wording used by the respondents. The first waves are characterized by greater use of “polar bears”, “typhoon”, “should”, and “natural catastrophe”,. We also observe that responses from the early waves are characterized by rarer use of terms like “plastic”, “temperature”, and “human-made”, compared to later waves.

Table 3. Keyness analysis by survey wave. Responses from waves 1 (2013) and 3 (2014) are compared to responses from waves 13 and 21 (2018 and 2021). Responses from wave 6 and 8 (2016 and 2017) are omitted. Relative frequency of a word in the two groups and percent difference in use (%DIFF).

Word	Translation	Rel. freq. wave 1 & 3	Rel. freq. wave 13 & 21	%DIFF
Isbjørner	polar bears	1.2	0.0	2713.2
Tyfon	typhoon	1.2	0.1	1306.6
Burd	should	1.8	0.4	330.2
Naturkatastrof	natural catastrophe	6.9	3.2	114.9
Klima	climate	16.7	25.8	-35.4
Menneskeskapt	human-made	12.9	21.8	-40.8
Oppvarm	warming	8.7	15.0	-41.9
Endring	change (noun)	8.7	17.3	-50.0
Kloden	globe	3.3	6.8	-51.0
Tørke	drought	2.7	6.4	-57.5
Endrer	change (verb)	1.9	4.6	-58.3
Temperaturen	temperature	0.8	2.8	-70.4
Plast	plastic	0.1	1.7	-96.0

Table 4 shows keyness analysis by degree of worry (binarized) about climate change. Respondents who express that they are worried about climate change use the words “does”, “serious”, “pollution”, and “extreme weather” more frequently than respondents that are not worried. Words used more frequently by those who are not worried include “nonsense”, “taxes”, “fuss”, “exaggerate(d)”, “cycle”, and “hysteria”.

Table 4. Keyness analysis by degree of worry (binarized). Very worried, worried and somewhat are coded as "worried". Not particularly worried and not worried are coded as "not worried". Relative frequency of a word in the two groups and percent difference in use (%DIFF).

Word	Translation	Rel. freq. worried	Rel. freq. not worried	%DIFF
Gjøres	does	1.7	0.2	644.7
Alvor	serious	2.0	0.4	432.8
Forurens	pollution	6.9	3.8	81.6
Ekstremvær	extreme weather	15.0	8.9	68.7
Klima	climate	20.7	28.4	-26.9
Klimaendring	climate change	7.9	12.5	-36.6
Tror	believe	2.9	5.6	-47.7
Naturlig	natural	5.8	12.1	-52.1
Avgift	fee	1.3	3.3	-58.7
Aktivitet	activity	1.0	2.7	-63.4
Alltid	always	2.0	5.5	-63.9
Istid	ice age	0.8	2.4	-66.0
Syklus	cycle	1.7	5.5	-68.6
Hysteri	hysteria	0.6	2.7	-78.5
Overdrev	exaggerate	0.2	1.3	-83.2
Overdrevet	exaggerated	0.3	1.6	-83.6
Mas	fuss	0.1	1.0	-86.8
Skatter	taxes	0.1	1.0	-86.8
Tull	nonsense	0.3	3.6	-91.5

Table 5 displays the keyness analysis of responses given by women versus men. The list of significant terms indicates that the language used in reaction to climate change is gendered. Women more often use words referring to changing weather and natural phenomena and seriousness, like "melts", "natural catastrophes", "scary" and "extreme weather". The top words refer to "sort", "garbage", women also tend to mention "plastic" more often than men. Among the words used more by men are "fees", "hysteria", and "money".

Table 5. Keyness analysis by gender. Women compared to men. Relative frequency of a word in the two groups and percent difference in use (%DIFF).

Word	Translation	Relative freq. Women	Relative freq. Men	%DIFF
Sorter	sort	0.5	0.0	n.a
Søppel	garbage	1.2	0.2	430.3
Skummelt	scary	2.3	0.5	336.7
Plast	plastic	1.5	0.4	244.7
Smelter	melts	8.9	3.1	184.8
Naturkatastrof	natural catastrophe	5.4	2.0	172.6
Dyr	animals	3.0	1.1	158.3
Skremmend	terrifying	2.9	1.1	157.8
Isen	ice	4.1	1.7	144.0
Storm	storm	2.4	1.0	136.2
Isbreer	glaciers	3.4	1.5	121.6
Polen	pole	3.4	1.6	112.4
Flom	flooding	5.8	3.0	97.3
Regn	rain	5.8	3.4	71.1
Forurens	pollution	7.6	4.8	56.9
Ekstremvær	extreme weather	15.9	10.5	51.4
Naturlig	natural	5.4	8.6	-36.9
Penger	money	0.8	2.2	-62.9
Liten	small	0.6	1.7	-66.3
Største	largest	1.0	2.9	-66.7
menneskeheten	humanity	0.5	1.6	-67.2
Vel	well	0.4	1.4	-68.4
Avgift	fee	0.7	2.8	-73.9
Hysteri	hysteria	0.4	1.7	-78.1
Solen	sun	0.1	0.8	-92.6
Primært	primarily	0.0	0.4	-100.0

3.2 Structural topic modeling

We ran a number of structural topic models, and selected one model with six topics for deeper analysis based on qualitative and quantitative assessments. Based on readings of the most representative words and responses, we gave the topics the following labels:

1. Weather
2. Policy responsibility/emotions
3. Attribution
4. Consequences for nature
5. Policy skepticism
6. Global warming

The *Weather* topic contains responses discussing changes in phenomena such as wind, precipitation, and seasons. Textual responses belonging to the *Policy responsibility/emotions*

topic often bring up global collective action, including by big companies, but not so much individual action. Emotions are often mentioned in this topic. The causes of climate change are prevalent under the *Attribution* topic. These are most frequently expressed as largely natural and cyclical in responses belonging to this topic, but some also emphasize human causation.

The *Consequences for nature* topic includes elements such as ice melt. The *Policy skepticism* topic frequently brings up problems with policies to reduce greenhouse gas emissions, often local and tied to transportation. Finally, the *Global warming* topic brings up a mixture of associations, with temperature change as an overall theme.

See Appendix 1 for the most representative responses by topic.

Table 6. Most representative terms by STM topic.

No.	Label	Seven most representative words (translation)
1	**Weather**	vind (wind), varmer (warmer), våtere (wetter), flom (flood), vintr (winter), snø (snow), regn (rain)
2	**Policy responsibility/ emotions**	redd (afraid), gjøres (do), land (land), tema (topic), folk (people), politikern (politicians), politisk (political)
3	**Attribution**	naturlig (natural), variasjon (variation), tiden (time), grad (degree), stor (big), skjedd (happened), utslipp (emissions)
4	**Consequences for nature**	endring (change), forurensn (pollution), endrer (changes), issmelt (ice melts), følge (consequence), forandr (change), ekstremvær (extreme weather)
5	**Policy skepticism**	største (biggest), negative (negative), gir (gives), går (goes), menneskelig (human), menneskeheten (humanity), syklus (cycle)
6	**Global warming**	flere (more), utfordr (challenges), hysteri (hysteria), ekstrem (extreme), mangel (lack), økend (increasing), katastrof (catastrophe)

Note. The seven most representative words by label are given, with representativity defined as a combination of frequency and exclusivity, see Roberts, Stewart, and Tingley (2014). Label names are suggested based on the most representative words as well as the most representative responses of each topic (see Appendix 1). The topics are ordered by aggregate topic prevalence, with the most prevalent topic (*Weather*) on top.

Explaining topic prevalence

What influences people's choice of topics? We here present relationships between topic prevalence for all or selected topics on the one hand, and age, worry about climate change, survey wave, and gender on the other. We start with age to evaluate our main research questions. We then examine correlations between topic prevalence and worry about climate change, in part to validate the chosen STM model with categorical opinion data. Third, we examine changes in topic prevalence over time, and finally, examine gender differences.

Topic prevalence and age

Figure 1 shows linear models for the two most prevalent topics over age. We find a positive correlation between age and prevalence of the *Weather* topic. Conversely, the younger the

respondents, the more likely they are to use words corresponding with the *Policy responsibility/emotions* topic. The remaining topics show only minor co-variation with age.

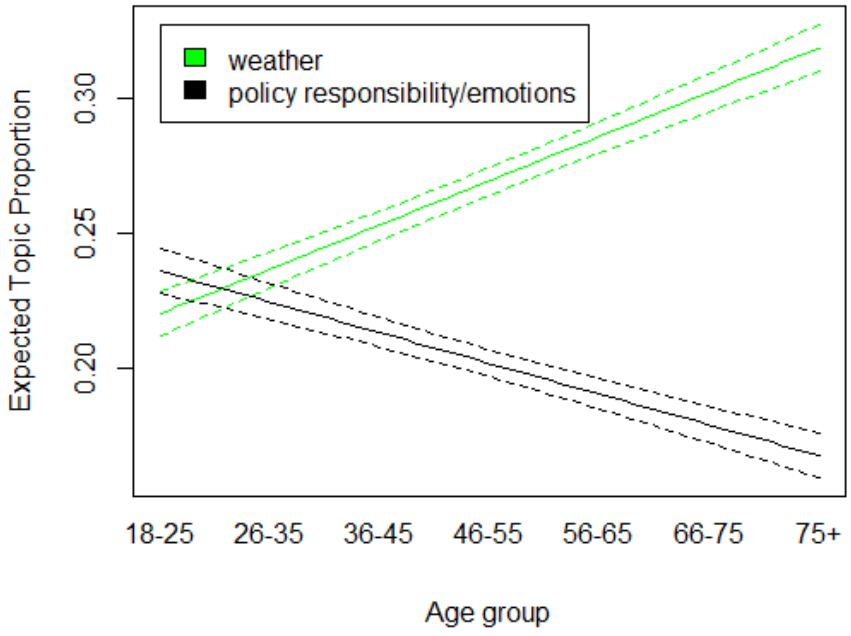

Figure 1: Predicted topic prevalence by age cohort, two selected topics. The figure is based on linear regression models with topic prevalence as the dependent variable and age cohort, gender, and survey wave number as explanatory variables. The inclusion of control variables does not substantially affect the results. Confidence intervals (95%) combine uncertainty from the regression model with uncertainty from the STM modeling process, see Roberts et al. (2014) for details.

To look more closely at the youngest group (born 1990 or later), we created figure 2 to contrast the youngest with the other age groups. This reinforces the finding that the youngest associate climate change less with the weather and more with policy responsibility. Specifically, the prevalence of the *Weather* topic is 0.19 for the young against 0.25 for the non-young. Furthermore, the young are more likely to use words related to the *Policy responsibility/emotions* topic, with a prevalence of 0.25, against 0.20 for those not in the youngest cohort. The young also tend to be more likely to address policy skepticism and global warming than the non-young. By contrast, the *Consequences for nature* topic shows slightly lower prevalence in the youngest cohort, although the significance of this difference may be called into doubt.

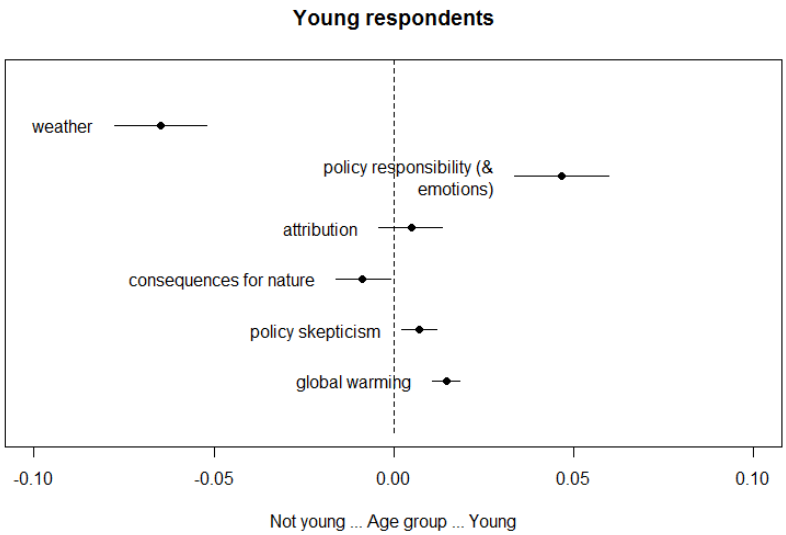

Figure 2: Contrasts in topic prevalence between young respondents (born 1990 or later) and remainder of the sample. The figure is based on linear regression models with topic prevalence as the dependent variable and age cohort, gender, and survey wave number as explanatory variables. The inclusion of control variables does not substantially affect the results. Confidence intervals (95%) combine uncertainty from the regression model with uncertainty from the STM modeling process, see Roberts et al. (2014) for details.

Topic prevalence and worry about climate change

The level of worry about climate change co-varies with the prevalence of several topics. However, the relationships are not generally linear. Figure 3 shows that the *Weather* topic is chosen least by those who are "very worried" as well as "not worried at all" about climate change – the extreme values of the scale. Those most likely to bring up the weather tend to select the worry levels between the extremes.

As regards the *Policy responsibility/emotions* topic, the most worried are significantly more likely to choose it, relative to all the other four levels of worry. The third most prevalent topic, *Attribution*, also shows a curvilinear relationship, with local peaks both at the highest and lowest level of worry. However, overall, the individuals with low levels of worry are more likely to use words linked with this topic.

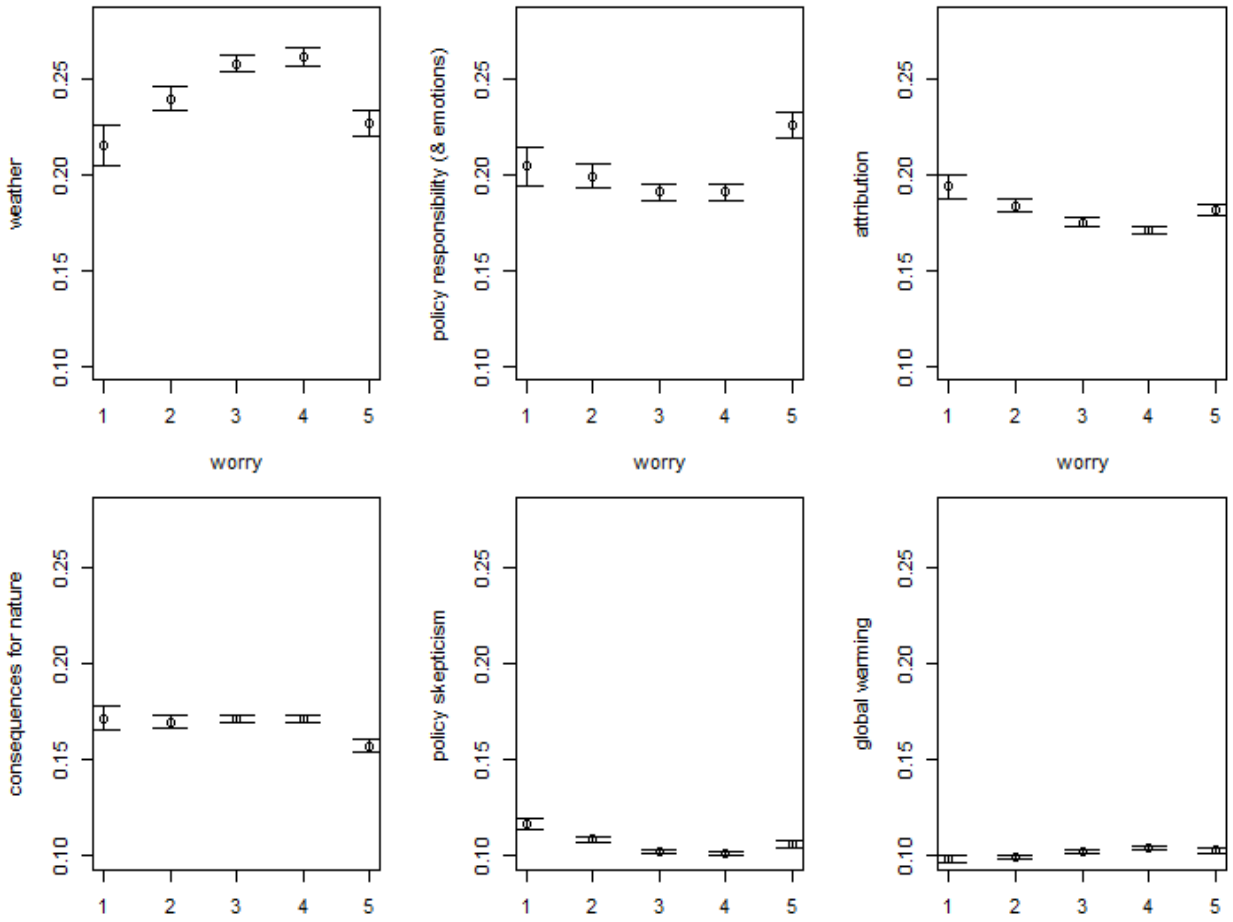

Figure 3: Topic prevalence over worry about climate change. Means and confidence intervals of 2 SE are given for each level of worry. Note that the uncertainty bounds do not reflect uncertainty in the STM estimation (see Roberts et al., 2014).

Topic prevalence over time

What is the development of topic prevalence over time? Figure 4 shows developments for each topic over six waves, for the entire sample. Most notably, the *Consequences for nature* topic increases markedly over time, whereas the *Policy responsibility* topic drops somewhat from a high in the third of the six waves, conducted in March 2016. We do not see marked differences between the youngest cohort and the remainder of the sample in the movement of topics over time.

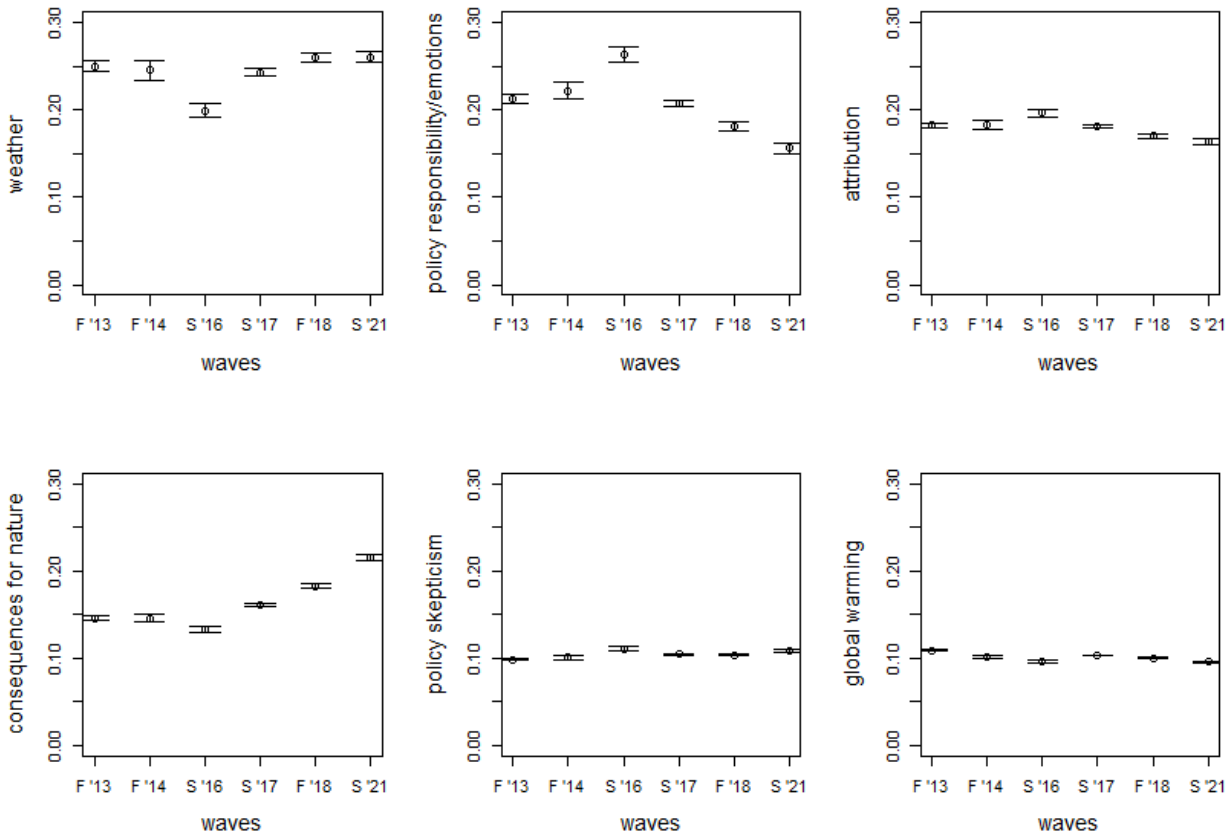

Figure 4: Topic prevalence over time. Means and confidence intervals of 2 SE are given for each survey wave. Note that the uncertainty bounds do not reflect uncertainty in the STM estimation (see Roberts et al., 2014).

Gender and topic prevalence

Our final explanatory variable is gender. As shown in Figure 5, we find that women are more likely than men to use words associated with the *Weather* and *Consequences for nature* topics, but less likely to use words related to *Policy responsibility/emotions*, *Attribution*, and *Policy skepticism*.

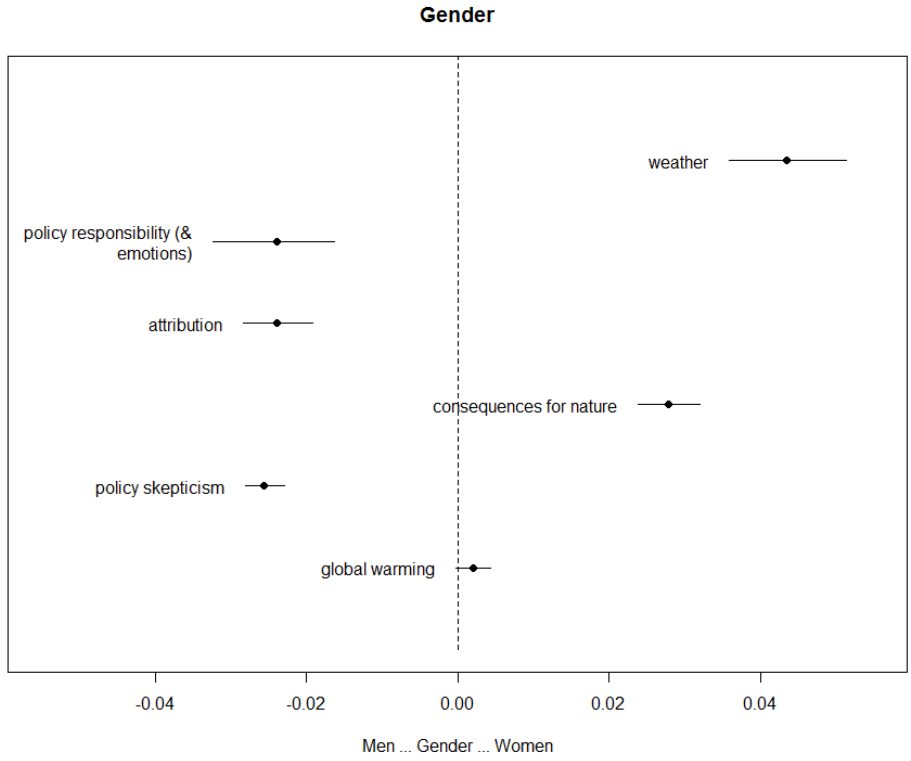

Figure 5: Topic prevalence by gender. The figure is based on linear regression models with topic prevalence as the dependent variable and age group, gender, and survey wave number as explanatory variables. Confidence intervals are at the 95% level and combine uncertainty from the regression model with uncertainty from the STM modeling process, see Roberts et al. (2014) for details.

4 Discussion and conclusion

We aim to map how people think about climate change, with a special focus on the young and on potential changes across time. To this end, we utilize open-ended responses that representative samples of the Norwegian public gave when being asked what first came to their minds when thinking about climate change. These responses were given in surveys as part of the Norwegian Citizen Panel from the end of 2013 to the summer of 2021.

In the following, we will discuss the results with respect to the research questions that we phrased in the introduction. First, how do people’s associations with climate change vary over time? Second, do people perceive climate change differently depending on their age, gender and worry about climate change?

When considering which words people mention most frequently overall -- across all survey waves and respondent groups -- we find that these most frequent associations refer to close descriptions of the meaning of climate change (e.g., climate, warming, global), to (extreme) weather, and to human causation. This pattern is consistent with the few other studies that have collected textual data in surveys to examine people’s associations with “climate change,” such as Lorenzoni et al.’s (2006) early investigation comparing US and British samples. Relatively universally, people’s first immediate associations tend to focus on global temperature rise,

predominantly weather-related impacts on nature, and the question of whether climate change is of anthropogenic origin, which is often addressed with climate skeptical connotations (Tvinnereim & Fløttum, 2015; Tvinnereim et al., 2017).

The keyness analysis reveals significant differences in language use between groups of respondents. For age cohorts, it shows that the young address more imperatives for action, the global nature of the problem, and the low frequency of terms that signify the weather-related impacts of climate change. This agrees with the results from the STM model analysis, where young people tend to use words related to the *Policy responsibility/emotions* topic, whereas older respondents bring up the *Weather* topic more. These results also show continuity from our 2015 study. The young is a relatively small group in our sample (966 out of 12,226 observations) and the group is underrepresented in the panel. This warrants us to be somewhat careful when we interpret our findings for the age cohorts.

A novelty of the current study is the opportunity to study changes in people's textual responses about climate change over time. Keyness analysis contrasting earlier and later waves shows both some interesting trends and some new aspects that have been introduced into Norwegian climate discourse. The iconic polar bear is relatively frequently referred to in the early waves but fades over time. In later waves, a new environmental problem, plastic, emerges. We observe a decrease in personal and normative terms in later waves ("I", "should") and an increase in terms that point to physical processes ("changes", "warming", "droughts", "human-made") and perhaps personal experiences with changing weather (in contrast to polar bears). The latest waves show a statistically stronger prevalence of words related to negative changes in nature, which suggests experiences with climate change in recent years. This observation fits the increasing prevalence of the *Consequences for nature* topic in the STM analysis. The differences between early and late waves remain almost the same if we include wave six as part of the early waves and wave eight with the last waves. This indicates that the observed temporal changes have been gradual and that the finding is robust.

The keyness analysis also provides a rich display of the words used by individuals who are less worried about climate change, as opposed to those that are more worried. The less worried and perhaps contrarian voices use words related to the cyclicity and natural causes of past climate changes phenomena ("natural", "cycle", "always") or consider the problem exaggerated or dismiss it as a nonsensical topic. They are also characterized by more frequent use of terms related to taxes and fees. Skepticism of pricing mechanisms to reduce emissions are likely behind this finding. This is in line, albeit weakly, with the connection made between climate skepticism and individual free-market ideological adherence posited by Lewandowsky, Oberauer, and Gignac (2013).

We also find that the language used to respond is gendered. Women more often employ terms that refer to natural phenomena and everyday actions like sorting garbage, but also to use adjectives like "scary" and "terrifying" more frequently than men do. Men, on the other hand, are more likely to use terms that refer to the cost of measures, such as taxes and fees, and also use abstract terms more ("humanity", "primarily"). Some of the words more characteristically used by men are also the same terms that characterise the language used by those that are not worried ("fee", "hysteria" and "natural").

The STM run chosen for this paper has some similarities with the one presented in Tvinnereim and Fløttum (2015). Notably, two topics, *Weather* and *Attribution* are quite similar (in 2015 the label was *Weather/Ice*). It is also notable that the *Weather* topic is the most prevalent in both studies. At the same time, the newer data set is about six times as large, and more topics have been found.

As regards predictors of topic prevalence, we find differences in associations across age groups. The young cohort associate climate change less with *Weather* and more with *Policy responsibility/emotions* compared to the older age cohorts. This resonates well with the finding by Tvinnereim and Fløttum (2015) that older respondents are more likely to express their thoughts about climate change in terms of weather, whereas young respondents tend to emphasize the social and human aspects of climate change. Moreover, the *Policy responsibility/emotions* topic is characterized by citizens using an emotional language, talking for example. about sadness and fear in relation to climate change. That the young cohort is overrepresented in this topic connects well with the findings reported earlier that young people are clearly more worried about climate change compared to the older age groups.

Moreover, the non-linear relationships between levels of worry and topic prevalences stand out. In particular, the fact that the *Attribution* topic has local peaks both among those “not at all worried” and those “very worried” suggests that the propensity to emphasize causes of climate change – human or natural – can derive both from a desire to dismiss climate change as a problem and from a real concern over what humans may be causing or could do. This finding departs somewhat from the earlier finding of a clear negative statistical correlation between the *Attribution* topic and worry about climate change (Tvinnereim and Fløttum, 2015). At the same time, that previous study did find from qualitative readings that the *Attribution* topic did not overwhelmingly express skepticism about human causation of climate change, but rather uncovered substantial amounts of ambiguity.

Across several topics, the group reporting the highest level of worry about climate change stands out. This group is clearly more likely to mention policy responsibility and attribution, while clearly less likely to mention the weather and consequences for nature, compared to the group selecting the second-highest level of worry. This suggests a qualitative difference – not just one of degree – for this group, and thus points out an avenue for future research.

A novelty of this study is the use of open-ended survey questions over time. Our analysis of topic proportions over time shows that *Consequences for nature* gains prevalence over time. This is a topic that is emphasized by women but not by the youngest cohort, nor by individuals who are the most worried about climate change. It is worth examining further what drives this increase over time.

Furthermore, the *Policy responsibility/emotions* topic spikes in the third survey (apparently in counterpoint to the *Weather* topic) before declining over time. The point at which *Policy responsibility/emotions* was briefly the largest in Wave 6, collected in March 2016. We conjecture that this emphasis on policy may have followed in the wake of the Paris Agreement, sealed a few months earlier. At the same time, the 2016 wave is the second smallest among the ones presented in this paper (N=1,016), potentially increasing uncertainty. Reference to the Paris

Agreement itself is also only made 29 times across the entire data set, with only one mention in the 2016 wave.[4] Closer analysis including qualitative readings from this wave may be necessary to explain the development of the *Policy responsibility/emotions* topic.

As regards gender and topic proportions, the findings largely agree with previous research (Tvinnereim & Fløttum, 2015), notably as women emphasize the *Weather* topic whereas men talk more about *Attribution*.

Future research should use the panel aspect of the data more. In the analysis presented here, we have assumed that each time a respondent provides a textual answer to a question, the answer represents a random draw from the population. However, there are good reasons to expect consistency within individuals over time, making this assumption perhaps overly strong, but also making new forms of investigation possible. Our data have over two thousand respondents who have given textual answers to our questions three times, and almost one thousand with four responses. This is a unique opportunity for following the development of individual opinions on climate change expressed in text over time.

Acknowledgments

This paper uses data from the Norwegian Citizen Panel waves 1, 3, 6, 8, 13, and 21 (Ivarsflaten et al., 2013 - 2021). The Norwegian Citizen Panel was financed by the University of Bergen (UiB) and Trond Mohn Foundation (TMS). Data collection was coordinated by UiB, implemented by Ideas2Evidence, and distributed by Sikt and UiB.

References

Benoit, K., Watanabe, K., Wang, H., Nulty, P., Obeng, A., Müller, S., & Matsuo, A. (2018). quanteda: An R package for the quantitative analysis of textual data. *Journal of Open Source Software*, *3*(30), 774.

Cornelis, I., Van Hiel, A., Roets, A., Kossowska, M., 2009. Age differences in conservatism: evidence on the mediating effects of personality and cognitive style. J. Pers. 77 (1), 51–88. https://doi.org/10.1111/j.1467-6494.2008.00538.x.

Gabrielatos, C. (2018). Keyness analysis: Nature, metrics and techniques. In C. Taylor & A. Marchi (Eds.), *Corpus Approaches to Discourse* (pp. 225-258): Routledge.

Gregersen, T. (2022). Bekymring for klimaendringer [Worry about climate change]. Retrieved from https://energiogklima.no/nyhet/bekymring-for-klimaendringer/

Grimmer, J., & Stewart, B. M. (2013). Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts. *Political Analysis*, *21*(3), 267-297. Retrieved from http://pan.oxfordjournals.org/content/21/3/267.abstract

Ivarsflaten, E. et al. waves 1, 3, 6, 8, 13, and 21 [Dataset], year 2013 - 2021. Data available from DIGSSCORE, UiB.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA Faked the Moon Landing—Therefore, (Climate) Science Is a Hoax: An Anatomy of the Motivated Rejection of Science. *Psychological Science*, *24*(5), 622-633. doi:10.1177/0956797612457686

[4] Mentions of "Paris" in the responses: 2013: 0, 2014: 0, 2016: 0, 2017: 11, 2018: 11, 2021: 7.

Lorenzoni, I., Leiserowitz, A., de Franca Doria, M., Poortinga, W., & Pidgeon, N. F. (2006). Cross-National Comparisons of Image Associations with “Global Warming” and “Climate Change” Among Laypeople in the United States of America and Great Britain. *Journal of Risk Research, 9*(3), 265-281.

McCright, A. M. (2010). The effects of gender on climate change knowledge and concern in the American public. *Population and Environment, 32*(1), 66-87.

Pickard, S. (2022). Young environmental activists and Do-It-Ourselves (DIO) politics: collective engagement, generational agency, efficacy, belonging and hope. *Journal of Youth Studies*, 1-21. doi:10.1080/13676261.2022.2046258

Poortinga, W., Whitmarsh, L., Steg, L., Böhm, G., & Fisher, S. (2019). Climate change perceptions and their individual-level determinants: A cross-European analysis. *Global Environmental Change, 55*, 25-35. doi:https://doi.org/10.1016/j.gloenvcha.2019.01.007

Roberts, M. E., Stewart, B. M., & Tingley, D. (2014). stm: R Package for Structural Topic Models.

Roberts, M. E., Stewart, B. M., Tingley, D., Lucas, C., Leder-Luis, J., Gadarian, S. K., . . . Rand, D. G. (2014). Structural Topic Models for Open-Ended Survey Responses. *American Journal of Political Science*(4), 1064-1082.

Skjervheim, Ø. et al. (2013 - 2021). Norwegian Citizen Panel methodology report, waves 1, 3, 6, 8, 13, and 21 [Produced by Ideas2Evidence].

Stoneman, P., Sturgis, P., & Allum, N. (2012). Exploring public discourses about emerging technologies through statistical clustering of open-ended survey questions. *Public Understanding of Science*.

Tvinnereim, E., & Fløttum, K. (2015). Explaining topic prevalence in open-ended survey questions about climate change. *Nature Climate Change, 5*, 744–748. doi:10.1038/nclimate2663

Tvinnereim, E., Fløttum, K., Gjerstad, Ø., Johannesson, M. P., & Nordø, Å. D. (2017). Citizens’ preferences for tackling climate change. Quantitative and qualitative analyses of their freely formulated solutions. *Global Environmental Change, 46*, 34-41. doi:http://dx.doi.org/10.1016/j.gloenvcha.2017.06.005

Tvinnereim, E., Lægreid, O. M., & Fløttum, K. (2020). Who cares about Norway's energy transition? A survey experiment about citizen associations and petroleum. *Energy research & social science, 62*, 101357. doi:https://doi.org/10.1016/j.erss.2019.101357

Appendix 1: Supplementary information

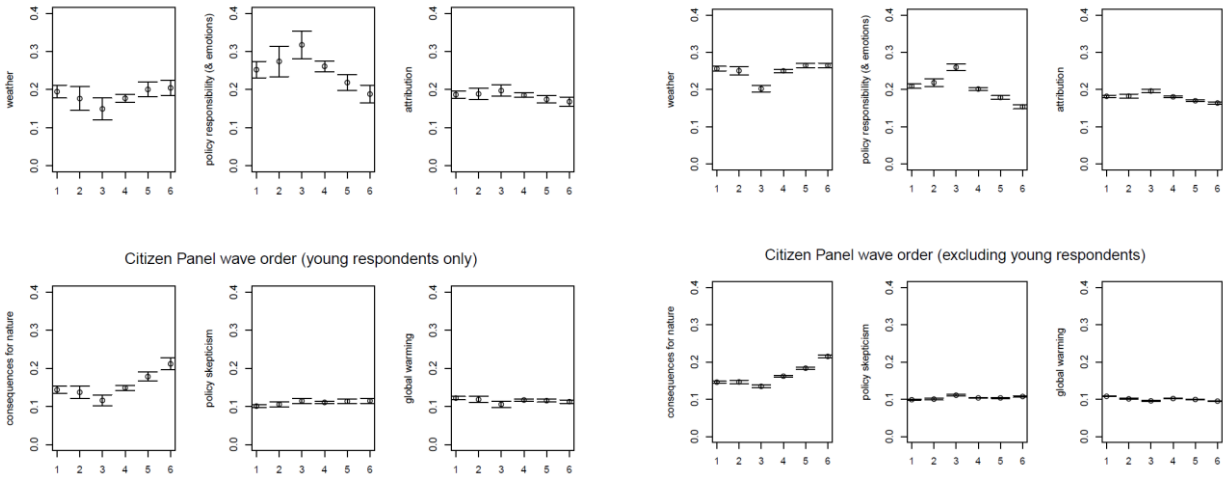

Figure A-1: Topic prevalence over time by age cohort.
"Young" respondents are those born 1990 or later (graph on the left). Topic prevalence for respondents born before 1990 is shown in the graph on the right.

Choice and description of STM model run

Initially, we ran two sets of STM runs: one with four topics and one with ten topics. We chose the number of four to be able to compare with earlier work using part of the data presented here (Tvinnereim and Fløttum, 2015) and chose the number ten to contrast with one that was substantially higher. Qualitative reading of the most representative responses and words for the resulting model runs indicated that a better number of topics would be found between four and ten.

We then ran STM models for all topic numbers from 3 to 30, using the `selectK` routine in R, and calculated semantic coherence and exclusivity for each. Along the coherence-exclusivity frontier, a handful of models dominated the others (6, 8, 9, 10). Among these, the model run with six topics had the highest level of coherence. We therefore selected this number of topics for further analyses see figure A-2.

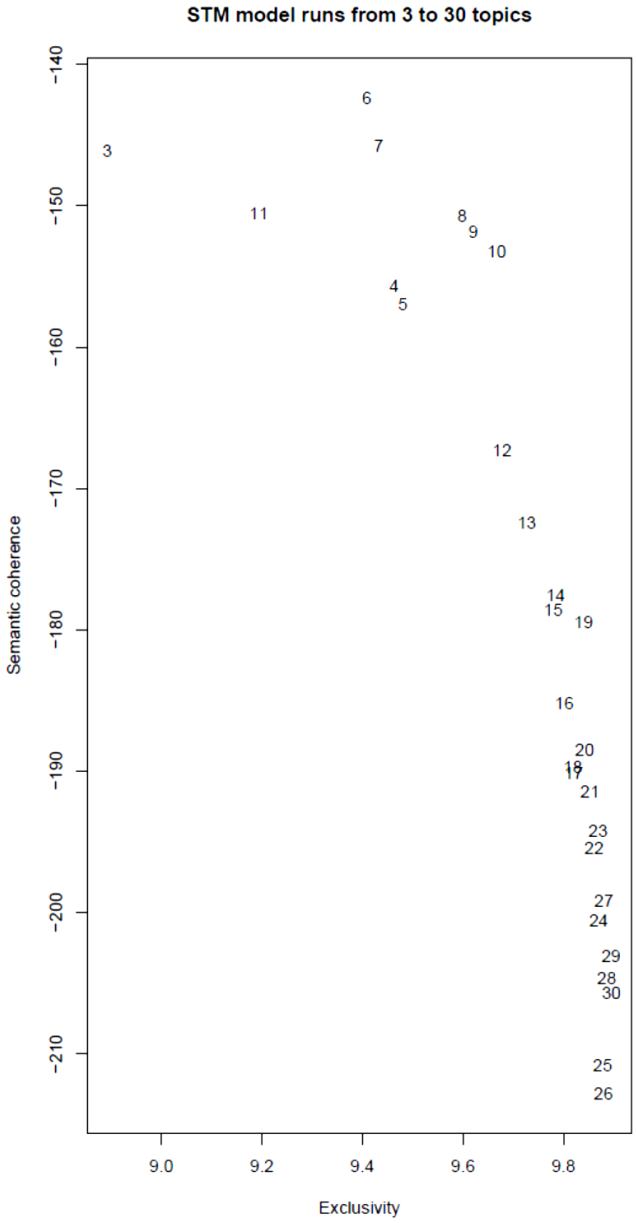

Figure A-2: Semantic coherence and exclusivity statistics for STM models with 3 to 30 topics

Table A-1: Most representative responses by topic

Topic # 1 most representative responses:
[1] "Ja okei tenker jeg først men det går for langt når dey gjelder bompenger. Fint med klimaendringer. Men bompengene går opp om mindre enn 5-10 år vil det gå opp i 200 kr. Det kan da ikke være \"dyrt\" for å komme inne i oslo? Hvorfor er det over flere bomstasjoner i oslo. Det er som om vi bor i roma igjen hvor oslo blir det sosiale og utenfor blir det som ikke er det sosiale. Som da romerne kom seg inn i torget for å snakke om politikk. Det samme blir oslo når oslo er blitt omringet av bom."
[2] "politikk og hvordan kunne tjene å få inn mest penger på div tiltak som evt skal fremme klima det er bevist at meste parten av de pengene som kommer inn via såkalte klimatiltak ikke går det nettopp klima men til andre såkalte goder i samfunnet"

[3] "At nå kommer enda en runde med ensidig skremsels-propaganda for å få folk til å godta høyere skatter og avgifter for å \"redde kloden\". Samt at vi skal godta et såkalt \"grønt skifte\" som får som resultat bl.a. at norsk natur blir pepret med monster-vindmøller til ingen nytte."
[4] "Globale temperaturer endrer seg i negativ retning (blir varmere). Dette kan være påvirket av naturlige sykluser samt menneskelig påvirkning. Det burde derfor tas på alvor."
[5] "Det er fakta, og mest sannsynlig den største utfordringen vi som mennesker står ovenfor. Størsteparten av jordas befolkning vil ikke bry seg før effektene av vår misshandling begynner å påvirke deres daglige liv og vaner. Dagens barn og fremtidige generasjoner vil lide for vår manglende handling."
[6] "Jeg tenker på at mange i Norge tror vi skal \"frelse\" verden ved å sette opp en masse vindmøller som ødelegger den fantastiske kystlinjen vi har. Når man kunne oppnådd bedre resultat ved å oppgradere gamle vannkraftanlegg til en mye lavere pris."
[7] "Klimaendringer er noe som vil eller har allerede påvirket oss alle. Den største trusselen er overbefolkning av kloden, men det vil ingen erkjenne ! Dette har negativ effekt med tanke på uttslipp og tilgang på mat/vann. Nå og i framtiden."

Topic # 2 most representative responses:
[1] "Global oppvarming Fornybar energi Havnivå Politisk handlingslammelse Global utfordring Lokale endringer i vær og vind (klimaendringer)"
[2] "Først og fremst global oppvarming. Men i forlengelsen av det ulike konsekvenser. Her tenker jeg på naturkatastrofer og hvordan klimaets har betydning for dyr og mennesker. Først og fremst utryddelsesfare for dyr, og mennesker i u-land hvor klimautviklingen kan få katastrofale konsekvenser steder hvor utgangspunktet for levestandarden allerede er svært lavt. Jeg tenker at klimaendring er et evig politisk diskusjonstema, som sjelden kommer noen vei."
[3] "Når jeg hører ordet klimaendring tenker jeg på global oppvarming generelt. Altså smelting av is på nordpolen og sør polen, bland annet."
[4] "Menneskeskapt. Fornybar energi. Mindre forbruk."
[5] "Tenker at det er veldig alarmerende, får mye katastrofe tanker. Tenker at vi må prøve å gjøre så mye vi kan for å stoppe global oppvarming før de er for seint."
[6] "Jeg tenker på global oppvarming, og de endringene som dette fører med seg, for eksempel tørke i Afrika, smelting av polene og at havnivået stiger."
[7] "Tilgjengelig forskning beviser at de er menneskeskapte, gjennom forbrenning av fossile brensler."

Topic # 3 most representative responses:
[1] "Global oppvarming: havnivået stiger, isen smelter, øyer forsvinner, mer uvær og flora og fauna endrer seg."
[2] "Global oppvarming, isbreer som smelter, endring i flora og fauna som følge av endringer i klima som igjen påvirker annen flora og fauna i kretsløpet osv."
[3] "Mer ekstremvær. Isbreene smelter. Påvirkning på både flora og fauna."
[4] "Endringer i klimaet kan skje naturlig, men tenker hovedsaklig på de endringer i klimaet som skjer pga. mennesker og vår aktivitet (utslipp, avskoging, osv.). Global oppvarming, stigning i havnivået, mer ekstremvær, dårligere luft, utryddelse av dyr og mulige klimaflyktninger."
[5] "Issmelting, forsøpling (særlig plast), ekstremvær, forurensning, gassutslipp."
[6] "gjennomsnittstemperaturen på jorden stiger, havisen og breer smelter, ørkenene vokser, endringer i økosystem, sult, ekstremvær"
[7] "Store endringer i klimaet som er menneskeskapt . Global oppvarming, ekstremvær, is smelting"
"
Topic # 4 most representative responses:
[1] "Jeg håper og ber at verdens land vil komme til enighet om en omfattende avtale. Det ER mulig å få gjort noe med dette problemet, men rike land må slutte å vere så grådige, og faktisk hjelpe fattige land, ved å invistere i klimavennlig industri og co2 minkende tiltak i fattige land så vel som i sine egne. Oljefondet bør trekkes ut av alt skittent og miljøfiendtlig, og heller brukes til investeringer i milkøvennlig teknologi, infrastruktur, osv i andre land. Videre bør FN opprette et fond, som alle rike land skal betale en kvote på f.eks 1% av BNP, eller hva som matamatisk trengs, til, der pengene går til fremmingen av miljøvennlige tiltak i u-land. Deretter bør rike land tvinges til å bruke ytterligere halvparten av hva verdi de spytter inn i fondet, ekstra, til tilsvarende prosjekter... eller bistand. Og på toppen av dette bør jo denne utslippsreduskjonsavtalen komme, med dette FN fondet, og I-lands ytterligere forpliktelser, som insentiver til U-land om å signere avtalen. Det er vi rike land som har skapt problemet, vi må ta mesteparten av ansvaret."
[2] "Det er skremmende, og det virker som ingen gjør noe for å minske problemet. Alt handler om makt og penger; vi blir, for eksempel, sjeldent informert om de store problemene (på miljøet) som

kommer med kvegdrift, ettersom kjøtt og meieri industrier er noen av de mektigste industriene i Norge, så er det viktig for dem at vi fortsetter å kjøpe så mye av deres produkter så mulig, siden da får de mer makt og mer penger. Selv av å vite at kvegdrift er svært skadelig for miljøet (fra sterke metangass som kyrne slipper ut), så velger de å se bort i fra konsekvensene, og fortsetter å engasjere folk til å spise kjøtt til hvert
måltid."
[3] "Norge er et lite land som ikke kan gjøre noe uten de store land gjør no, som å kaste peng ut vinduet.det hjelper lite om ikke de som virkelig forurenser gjør no.så gidd ikke de hvorfor vi.?."
[4] "Føler en viss redsel, spesielt med tanke på kommende generasjoner, er redd vi ikke gjør nok for å redde kloden. Føler at vi som vanlige borgere kan gjøre noe, men monner det noe i den store
sammenheng?"
[5] "Uten håp. Alt styres av makt og penger. De som har tilgang til naturressurser både offentlige og private vil aldri slippe taket hvis muligheten for inntekt er
der"
[6] "Altfor mye mas omkring klimaendring. Begrenset hva lille Norge kan bidra med. Hva med USA, Kina og de virkelig store ? Hører ikke så mye om hva de gjør for å endre
utslippene:)"
[7] "Problemet vil ikke finne sin løsning med den type virkemidler vi diskuterer i dag. Befolkningsreduksjon, slutt på proteinproduksjon og andre viktige virkemidler skyves ut til fordem for
symbolpolitikk."

Topic # 5 most representative responses:
[1] "Det er dokumentert naturlige klimaendringer I klodens historie gjennom millioner av år, med varmetider og istider. Klimapanelet har dokumentert at det har skjedd betydelige klimaendringer etter at industialiseringen tok fart og med den økte bruken av fossil brensel. Paris avtalen var et gjennombrudd for global innsats for å begrense utslipp som bidrar til at klimaendringer
utfases"
[2] "Klimaendringer he alltid eksistert. Ogso for tusen år sida. me he hatt istid og smeltetid. Sån vil de alltid vær. so ditta ekje nåke og skrike for. kor vidt klimaendringe he me co2 utslepp å gjer, det veit ei ikkje. Men tvila på att de e avgjerandes"
[3] "Et alvorlig problem som må tas tak i. Både når det kommer til en god utnyttelse av resurser på lokalt nivå, som kollektivt eller kildesortering. Og når det kommer til klimagasser på et globalt nivå, som syrning av havet og utslipp av drivhusgasser."
[4] "Naturlige variasjoner i jordens klima - basert på variasjoner i solens prosesser og på naturlige prosesser på jorden, inkludert menneskelig
aktivitet."
[5] "Jeg mener at klimaendringer har det alltid vært, det var varmt i romertiden, i vikingetiden og ganske varmt nå. i mellomtiden var det kaldere perioder som i lille istiden på 14 - 15 hundtetallet. Det er ikke noe spesiellt som hender nå."
[6] "Klimaendringer har skjedd i tusenvis av år tilbake. Der er det sykluser vi ikke kan gjøre stort med. Men dermed mener jeg i k k e at vi skal gi blaffen i miljøforsøpling ! Der m å vi skjerpe oss globalt
!!"
[7] "Historien viser at det har vært klimaendringer opp gjennom tidene. Vi påvirker den i større eller mindre grad. Sannsynligvis er forskningen for sterkt opptatt av en årsak og ikke hva vi kan ha som motvekt ... vi må finne balansepunktet"

Topic # 6 most representative responses:
[1] "Mye nedbør i form av regn, sterk vind, milde vintre, oversvømmelser, ras."
[2] "Mer ustabilt klima både på temperatur, nedbør og vind. Vil føre til mer skader på hus og byggverk"
[3] "Høyere temperaturer, mer nedbør, høyere vannstand, nedsmelting av isbreer og polområdene"
[4] "En fare for jorda En fare for freden på jorda En fare for mennesker og dyr En fare for demokratiske styresett En fare for sivilisasjoner"
[5] "Mer intens nedbør. For østlandet, mildere vintre med mer nedbør, varmere og tørrere somre. Smelting av innlandsis. Heving av havnivå."
[6] "Mer nedbør. Oversvømmelse, ubeboelige områder, flom, tap av inntektsgrunnlag og livsgrunnlag. Ubehagelig klima."
[7] "Menneskeskapte forandringer på klimaet, mildere vintre, mer nedbør og mer vind."